

S VENKATESH

+91 6300645276

mailto: venkateshsurvi14@gmail.com

OBJECTIVE

Experienced Data Engineer with 5.11 years in the IT industry, in designing, developing, and implementing ETL processes and data pipelines. Proficient in SQL, and Pyspark. Experienced in cloud computing platforms like Azure and AWS . Expert in data lake, data warehousing, and data analytics with a track record of delivering high-quality data solutions.

- Good Knowledge on Data Integration Solutions using Microsoft Azure, ADF, ADLS, Storage Accounts, AWS S3, Redshift ,Azure Data brick, pyspark and Informatica Power center.
- Experienced in interacting with Clients, Business Analysts, UAT Users and Developers.
- Strong business understanding of verticals like Insurance, Banking, health care, e-Commerce and retail.
- Ability to meet deadlines, handles multiple tasks, decisive with strong leadership qualities, flexible in work schedules, and possesses good communication skills.

EDUCATION

Master of Technology : Jawaharlal Nehru Technological University, Hyderabad

SKILLS

Languages	Pyspark, Shell Scripting
ETL Tools	Informatica Power Center, SSIS
Cloud Platform	Azure, AWS, Databricks
Database	SQL Server, Oracle, DB2
Scheduler	Redwood RMJ,Control-M, Airflow
Control System and Documentation	SVN, Git, Jira, Confluence

CERTIFICATIONS

DP-203 : Azure Data Engineer Associate

EXPERIENCE

PROJECT #1:

July 2021 – May 2023

Primus Global Technologies with Cognizant

Hyderabad,TS

Client: DLG - UK

Role: Azure Data Engineer

Direct Line Group Insurance plc is a British insurance company based in the United Kingdom. DLG UK Motor Insurance Project where the data flows comes from Guide wire (Policy Center, Claim Center and Billing Center) to Data Mart and Data lake.

Key Responsibility:

- Worked on Azure Cloud tools like ADLS, ADF, Azure Database, Triggers, CI/CD, Logic apps, Azure Key vault, Data bricks Pyspark and Azure synapse.
- Extensive experience on creating ADF Pipelines, Data flows, Storage account, Azure IR, optimization, query optimization and data bricks code optimization, Scheduling triggers and monitoring pipelines.

- Created Medium and complex initial & incremental load Pipelines to move data from different heterogeneous sources like on-premise SQL Server, Oracle server, Flat files etc into Azure Data Lake and then into Azure SQLDB and Azure Data bricks.
- Created dynamic pipelines and datasets as re-usable component to carry out data transformations and data movement.
- Hands on experience in creation of Pipelines with different activities like Copy Data, Lookup, For-Each, Pipeline, set variable, Wait and Data flows.
- Hands on experience with Data Flow activity and transformations like Join, Derived column, Lookup, Aggregate, Pivot, alter row, exists, conditional split, etc.
- Hands on experience with temporal tables and ADF pipelines to carry out SCD Type1 and Type2 loadings.
- Good knowledge and working experience on Indexes, CTE's, Temp Tables and views.
- Experience on diagnosing and resolving the pipeline failures and issues.
- Experience in performance tuning, tuned long running pipelines and queries.
- Worked on scheduling pipelines and cost optimization.
- Experienced in creating CI/CD pipelines to deploy ADF pipelines to higher environments.
- Ability to meet deadlines, handles multiple tasks, decisive with strong leadership qualities, flexible in work schedules, and possesses good communication skills.
- Working in Agile environment and participate in different phases of the project cycle like production and Technical support (Deployment phase).
- Worked on three layered (Brownze,Silver,Gold) Data lake architecture and Multiple Transformation using Pyspark.

PROJECT #2 :

Netfinity Technologies Pvt Ltd with Accenture

Client: BT - UK

Role: Azure Data Engineer

May 2019 – July 2021

Bengaluru, KA

BT Group Plc, trading as BT, is a multinational telecommunications Services Company headquartered in London, United Kingdom. we extract the data from Flat files and Oracle database which contains customer pre-paid and post-paid information in this system.

Key Responsibility:

- Have good experience working with Azure BLOB and data lake storage.
- Experience in developing Spark applications using Spark-SQL in Databricks for data extraction, transformation, and aggregation from multiple file formats for Analyzing & transforming the data to uncover insights into the customer.
- Extract Transform and Load data from sources Systems to Azure Data Storage services using a combination of Azure Data factory. Data ingestion to one or more Azure services (Azure Data Lake, Azure Storage, Azure SQL,Azure DW) and processing the data in Azure Synapse.
- Created Pipelines in ADF using Linked Services / Datasets / Pipeline / to Extract, Transform and load data from different sources like Azure SQL, Blob storage, Azure SQL Datawarehouse, write-back tool and backwards.
- Delivering new and complex high-quality solutions for Users in response to varying business requirements.
- Process structured, semi-structured and unstructured data sets from multiple platforms to identify & assess key takeaways.
- Engineer modern data pipelines to collect, organize, and process data near real time using PySpark and Python.
- Experienced in creating CI/CD pipelines to deploy ADF pipelines to higher environments.

PROJECT #3 :

Netfinity Technologies Pvt Ltd with Accenture

July 2017 - May2019

Bengaluru, KA

Client: WISCONSIN HEALTHCARE SYSTEMS - MADISON, WI

Role: Informatica Developer

The goal of the project was to develop an enterprise data warehouse to implement a centralized database that collects, organizes and stores data from different operational data sources to provide a single source of integrated and historical data for the purpose of reporting, analysis and decision support and also to improve the client services. The data is integrated from sources like flat files and loaded into a central data warehouse using transformations.

Key Responsibility:

- Worked closely with the business analyst and Data warehouse architect to understand the source data and need of the Warehouse.
- Involved in designing of star schema based data model with dimensions and facts.
- Created mappings with heterogeneous sources like flat files, MS Access, Oracle databases and created targets in data warehouse using Informatica PowerCenter.
- Created Source to Target mapping from design and requirement documents.
- Built reusable transformations for recurring business logics using mapplets and used them in multiple mappings.
- Designed and developed mappings using Source qualifier, Aggregator, Joiner, Lookup, Router, Sequence generator, Expression, Filter and Rank transformations.
- Designed and developed test cases for unit and system testing.
- Created parameters and variables for incremental data loading effectively using Informatica workflow manager.
- Created event raise, event-wait tasks for maintaining dependencies between workflows.
- Efficiently interacted with UAT team and supported system testing and fixed bugs in mappings during QA phase.
- Optimized or tuned mapping sessions during load testing in QA environment.
- Created different transformations for loading the data into targets like Joiner transformation, Update Strategy, Lookup transformation, Expressions, Filter and Router.
- Worked extensively with mappings using expressions, aggregators, filters, lookup.
- Designed, developed, and tested various enhancements.
- Responsible for production support for the ongoing jobs they run every day.
- Hands on experience in Redwood RunMyJobs is the strategic batch scheduler.
- Delivered the assigned work within specified timelines with high quality.